

United States District Court  
For the Northern District of California

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

IN THE UNITED STATES DISTRICT COURT  
FOR THE NORTHERN DISTRICT OF CALIFORNIA

ORACLE CORPORATION, et al.,

Plaintiffs,

v.

SAP AG, et al.,

Defendants.

No. C-07-01658 PJH (EDL)

**ORDER REGARDING SCOPE OF  
DISCOVERY OF ELECTRONICALLY  
STORED INFORMATION, INCLUDING  
LIMITS ON NUMBER OF CUSTODIANS  
TO BE SEARCHED AND SAMPLING**

On July 1, 2008, the Court held a discovery conference in this case brought under the Copyright Act, 17 U.S.C. §§ 101, et seq., and the Computer Fraud and Abuse Act, 18 U.S.C. §§ 1030, et seq. in which Plaintiffs allege that Defendant TomorrowNow, which has been acquired by Defendant SAP, improperly downloaded software and support materials on behalf of its customers in violation of those customers' licenses with Oracle. For the reasons stated at the conference and set forth herein, the Court issues the following Order.

**Custodian limits**

In addition to other document discovery, such as from centralized repositories, the parties agree that discovery of custodians should be focused on a subset of each party's potentially relevant custodians and agree that the opposing party may select who to designate as custodian and in which order the searches and production will be done. However, the parties disagree on the number of custodians each should search. Plaintiffs have reduced the number of proposed custodians of Defendants for Defendants to search to a total of 165, while Defendants continue to propose that they search 115. The issue is one of proportionality under Federal Rule of Civil Procedure

1 26(b)(2)(C)(iii). Plaintiffs argue that discovery from 165 custodians is necessary and proportionate  
2 to the needs of the case, while Defendants respond, inter alia, that because they are only capable of  
3 producing documents from an average of ten custodians a month, and the average cost per custodian  
4 is about \$100,000, discovery of 165 custodians would take more than one year, past the discovery  
5 cutoff, and cost approximately \$16.5 million.

6 In determining proportionality, that is, whether the burden or expense of the proposed  
7 discovery outweighs the likely benefit, the Court considers the following factors: “the needs of the  
8 case, the amount in controversy, the parties’ resources, the importance of the issues at stake in the  
9 action, and the importance of the discovery in resolving the issues.” See Fed. R. Civ. P.  
10 26(b)(2)(C)(iii). Plaintiffs argue that millions of dollars spent on this one portion of discovery is  
11 justified given that Defendants allegedly engaged in serious intellectual property violations through  
12 extensive illegal downloads, creation of multiple infringing software environments, and theft of  
13 “many hundreds” of customers, such that it “appears” that damages are at least “several hundreds of  
14 millions of dollars and likely are at least a billion dollars.” Jt. Discovery Conference Statement  
15 (docket number 102) at 3. Further, Defendants, like Plaintiffs, are a large company with  
16 considerable resources.

17 Defendants counter that even while employing in excess of fifty contract attorneys and a  
18 software vendor in addition to litigation counsel, they have only been able to produce documents at  
19 the monthly rate of ten average-size custodians based on Defendants’ ability to review 500,000  
20 documents per month (i.e., on average, each custodian had 48,000 documents in his or her files), at  
21 the average cost of \$100,000 per custodian. Accordingly, discovery of 165 custodians as Plaintiffs  
22 urge would cost the staggering sum of \$16.5 million – on top of other discovery costs from searches  
23 of centralized repositories and targeted searches, not to mention lay and expert depositions and  
24 interrogatories. Moreover, the amount of time such an extensive production would take at the  
25 current rate would run afoul of the firm discovery cut-off date of June 19, 2009 and trial date of  
26 February 8, 2010 set by Judge Hamilton. While Plaintiffs argue that Defendants have created some  
27 of their own burden by including irrelevant material, they do not argue that Defendants have  
28 understaffed their efforts or that the costs are inflated or out of line with the cost of similar efforts,

1 provided that Defendants present an adequate evidentiary basis to support the average cost of  
2 \$100,000 as they have agreed to do no later than July 8, 2008. Further, Defendants point out that  
3 Plaintiffs have not provided any real analysis of their damages, and argue that rather than “many  
4 hundreds” of customers, only roughly seventy converted from Oracle to TomorrowNow and then  
5 SAP. In addition, although Plaintiffs believe that extensive custodian discovery is imperative, the  
6 Court notes that Plaintiffs did not propose shifting any part of these enormous costs to them.

7 While this case is undoubtedly important to the parties and implicates the public policy  
8 against theft of intellectual property, and may perhaps involve millions of dollars, those factors do  
9 not make the needs of the case so extraordinary that this Court can justify forcing Defendants to  
10 spend an additional \$5 million (subject to verification as promised regarding the actual cost to date)  
11 on custodian searches alone, beyond the already costly search of 115 custodians to which they have  
12 agreed (at an expected cost of \$11.5 million). The Court is mindful that the cost of discovery is only  
13 one consideration among many in determining the proportionality of discovery. However, the Court  
14 will not condone forcing Defendants to expend \$16.5 million on custodial searches for just one  
15 aspect of what will undoubtedly be extensive discovery in this case.

16 Most fundamentally, this Court must construe and administer the discovery rules in  
17 accordance with the wise mandate of Rule 1: “the just, speedy, and inexpensive determination of  
18 every action and proceeding.” Most major document-intensive cases in federal court can be said to  
19 involve important public policies, such as the integrity of our securities markets, the viability of  
20 competition and the protection of intellectual property, and not infrequently expose parties to millions  
21 of dollars in potential damages. Many of the parties (although by no means all), like Defendants  
22 here, have large resources. However, if they have to expend many millions of dollars on discovery in  
23 each lawsuit that fits this profile because of the explosion in the amount and types of electronically  
24 stored information, to recast Calvin Coolidge’s saying, the business of America (and other nations)  
25 will no longer be business. Instead, businesses will be mired in exorbitantly costly litigation where  
26 the process of discovery becomes so onerous and prohibitively expensive that it no longer facilitates  
27 resolution on the merits as a means to an end, but dominates and derails the litigation process.  
28 Indeed, if the scope of discovery in this case is not contained and managed in an efficient and cost-

1 effective way, it will run afoul of the deadlines set by the trial judge and produce an avalanche of  
2 information that could not be presented to a jury in a manner and time frame that it could digest.

3 On balance, after consideration of all of the factors, the Court concludes that custodian  
4 discovery should be limited to 120 custodians, subject to Defendants' proof of the estimated costs of  
5 custodian discovery. Defendants assure the Court, and Plaintiffs recognize, that such verification will  
6 be filed no later than July 8, 2008. As stated by the Court at the conference, Plaintiffs were to have  
7 notified Defendants by e-mail by July 2, 2008 regarding the form of proof that Plaintiffs would  
8 consider sufficient. The Court declines Plaintiffs' invitation to impose non-mutual custodian limits  
9 on the parties, specifically, to order Plaintiffs to produce discovery from one-third to one-half of  
10 Defendants' total number of custodians. Judge Hamilton imposed equal discovery limits on the  
11 parties and Plaintiffs have failed to make an adequate showing why custodian discovery should be  
12 unequal. As Plaintiffs themselves point out, they contemplate seeking millions of dollars if not more  
13 in damages, and Defendants may need to explore that area proportionately. In so ruling, however, the  
14 Court does not suggest that Defendants necessarily should use their full allotment of custodians,  
15 which is a ceiling, not a floor. The parties may reserve any number of the 120 custodians for later  
16 discovery. The Court does not expect to increase the number of custodians absent a showing of  
17 extraordinarily good cause.

### 18 **Sampling and extrapolation**

19 The Court has urged the parties to consider sampling of the voluminous evidence, to provide a  
20 basis for the parties' experts to extrapolate the extent of illegal activity and associated damages, in  
21 lieu of a likely futile and undoubtedly exorbitantly costly and labor-intensive attempt to produce and  
22 review all of it. See, e.g., Fed. Jud. Ctr., Manual for Complex Litig. (Fourth), § 11.493 (2004)  
23 (“Acceptable sampling techniques, in lieu of discovery and presentation of voluminous data from the  
24 entire population, can save substantial time and expense, and in some cases provide the only  
25 practicable means to collect and present relevant data.”). Plaintiffs propose applying this approach to  
26 the Defendants' use of software environments to create allegedly infringing derivative works and  
27 distribute such works. There may be other topics that would also lend themselves to this approach.  
28 The parties agree that this approach has merit, but need to agree on a detailed protocol with the

1 assistance of their experts. Accordingly, no later than July 18, 2008, the parties shall file a proposal  
2 for extrapolation, at a minimum as to software environments. The parties should consider including  
3 as part of any proposed protocol a provision that they may not attack each other's experts'  
4 conclusions on the grounds that the conclusions were reached based on an agreed-upon sampling of  
5 the evidence.

6 **Damages causation evidence**

7 Plaintiffs argue that they have not received comparable damages causation evidence from  
8 Defendants, such as, a summary report as to why Plaintiffs' customers did or did not move to  
9 TomorrowNow, or documentation of any of the claimed Safe Passage deals. At the hearing,  
10 Defendants stated that they would produce exemplars of win/loss reports, customer surveys, and the  
11 electronic customer relationship management database which they believe are comparable to what  
12 Plaintiffs produced. Defendants shall provide the exemplars no later than July 11, 2008. Defendants  
13 also agreed to provide a list of the roughly seventy customers that moved to TomorrowNow and/or  
14 SAP; the list should include the more clearcut instances of transfer as well as the less clear transfers,  
15 annotated accordingly, and production should begin promptly with the more clearcut cases.  
16 Defendants shall provide this list no later than July 18, 2008.

17 **Search terms**

18 Plaintiffs contend that they have formulated a list of search terms that hits one hundred  
19 percent of the responsive documents, while limiting the volume of documents that must then be  
20 manually reviewed for privilege and the like, before production. Defendants have not yet hit one  
21 hundred percent with an acceptable attendant decrease in the number of documents that must be  
22 manually reviewed, but are continuing to refine their efforts. The parties shall resolve this issue no  
23 later than July 11, 2008, and should be guided by the Court's comments at the hearing that it is not  
24 necessary or likely practical for Defendants to establish a search protocol that ensures reaching one  
25 hundred percent of responsive documents while also providing meaningful cost savings by reducing  
26 the volume of documents that must be manually reviewed.

27 **Targeted searches**

28 The parties agree that targeted searches should be used to conduct reasonably well-defined

1 searches for information in likely sources other than the files of the designated custodians. Targeted  
2 searches can include searches of centralized repositories on specific topics or of employees' files  
3 based on asking them to identify relevant documents on particular topics. The parties also agree that  
4 targeted searches should be limited to ten per side. Further, the parties agreed at the hearing that any  
5 limited targeted searches would not count against the custodian limit. Accordingly, no later than July  
6 18, 2008, the parties shall propose to the Court a plan for the timing of the targeted searches.

7 **Board meeting minutes**

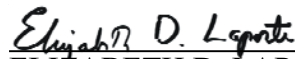
8 At the hearing, Defendants proposed treating Board meeting minutes as very highly  
9 confidential and redacting them for irrelevant information. No later than July 18, 2008, the parties  
10 shall propose a plan for handling discovery of Board meeting minutes.

11 **Further discovery conferences**

12 Further discovery conferences are scheduled for July 24, 2008 at 2:30 p.m. and August 28,  
13 2008 at 10:00 a.m. No later than one week before the discovery conference, the parties shall submit a  
14 joint discovery conference statement setting out in a non-argumentative fashion the substance of any  
15 discovery disputes, organized issue by issue, that is, setting forth each parties' position on each issue  
16 and proposed reasonable resolution before moving on to the next issue.

17 **IT IS SO ORDERED.**

18 Dated: July 3, 2008

19   
20 ELIZABETH D. LAPORTE  
21 United States Magistrate Judge  
22  
23  
24  
25  
26  
27  
28